A Deception-Driven AI Framework for Detecting Cyberattacks in Mobile Financial Applications Using GANs, Reinforcement Learning, and SVM

Chidiebere, Abigail Nnenna, Abiodun Esther Omolara, Mohammed Hammawa
Department of Computer Science, University of Abuja, Gwagwalada, Nigeria

Abstract
Mobile financial applications have become essential infrastructures for banking, digital payments, and financial services. However, their rapid adoption has significantly expanded the cyberattack surface of financial institutions, exposing users and service providers to increasingly sophisticated threats such as credential theft, account takeover, malware injection, and automated fraud attacks. Conventional rule-based and signature-driven security mechanisms often struggle to detect these evolving threats because they rely on previously known attack patterns and static detection rules. This study proposes a deception-driven artificial intelligence framework for detecting cyberattacks targeting mobile financial applications. The framework integrates Generative Adversarial Networks (GANs) to generate realistic synthetic attack scenarios, Reinforcement Learning (RL) to dynamically adapt deception strategies, and Support Vector Machine (SVM) classification to identify malicious interactions. The proposed system operates within a layered deception architecture that deploys honey accounts, honey transactions, and authentication traps to lure attackers into controlled environments while simultaneously collecting behavioral intelligence for threat analysis and detection. Experimental evaluation using benchmark cybersecurity datasets (NSL-KDD and CICIDS2017) augmented with GAN-generated attack samples demonstrates the effectiveness of the proposed approach. The integrated GAN–SVM detection model achieved 96.5% classification accuracy with a false positive rate of 1.8%, outperforming reinforcement learning based detection (93.2% accuracy) and standalone SVM classification (88.9% accuracy). Additional ROC–AUC analysis and ablation experiments further confirm the advantages of combining generative attack simulation with adaptive deception strategies. The results indicate that integrating deception technologies with artificial intelligence can significantly improve cyber threat detection performance while maintaining low false alarm rates. The proposed framework provides a scalable and proactive cybersecurity approach capable of enhancing the resilience of financial technology systems against evolving cyber threats.

Keywords:
Deception technology; honeypots; honey transactions; mobile banking security; generative adversarial networks; reinforcement learning; support vector machines; anomaly detection

1. Introduction

The rapid adoption of mobile financial applications has transformed the modern financial ecosystem by enabling users to perform banking transactions, manage digital wallets, and access financial services through smartphones and connected devices. Mobile banking platforms and financial technology applications provide convenience, accessibility, and improved financial inclusion for millions of users worldwide. However, this rapid digital transformation has also expanded the attack surface of financial institutions. Mobile financial applications operate within highly distributed environments that involve mobile devices, cloud services, application programming interfaces, and third-party service providers. These interconnected infrastructures introduce multiple security vulnerabilities that can be exploited by cyber adversaries seeking financial gain or unauthorized access to sensitive financial data (Almalki & Alotaibi, 2023).

Cyber-attacks targeting digital financial systems have increased in both frequency and sophistication in recent years. Threat actors frequently employ techniques such as credential stuffing, phishing campaigns, account takeover attacks, malware injection, and automated fraud activities to compromise financial platforms. Many of these attacks are adaptive and continuously evolve to bypass traditional security controls. Conventional defense mechanisms such as rule-based detection systems and signature driven intrusion detection often rely on previously known attack patterns. As a result, they struggle to detect emerging threats that do not match existing signatures or predefined rules. This limitation highlights the need for more intelligent and adaptive cyber defense mechanisms capable of anticipating attacker behavior and responding dynamically to evolving threats (Behl & Behl, 2024).

Deception technologies have emerged as a promising approach for strengthening cyber defense strategies in complex digital environments. Unlike traditional security systems that primarily focus on blocking attacks, deception-based defense mechanisms intentionally

Manuscript received March 5, 2026
Manuscript revised March 20, 2026
https://doi.org/10.22937/IJCSNS.2026.26.3.3

deploy decoy resources such as honeypots, honeytokens, and simulated transactions to mislead attackers. These deceptive assets create controlled environments that attract adversaries and encourage them to interact with fake resources that appear legitimate. Such interactions generate high quality telemetry that can reveal attacker intentions, tactics, and behavioral patterns. By capturing intelligence from these interactions, deception systems can detect malicious activity at early stages and provide valuable insights for improving defensive strategies (Fraunholz & Schotten, 2022).

Recent advances in artificial intelligence have further enhanced the capabilities of cyber defense systems. Machine learning techniques have been widely applied to analyze network traffic, system logs, and user behavior patterns for detecting anomalies and suspicious activities. Natural language processing, deep learning models, and statistical learning methods have demonstrated strong performance in identifying cyber threats across different domains. In particular, generative models and reinforcement learning techniques provide promising opportunities for developing adaptive and proactive cybersecurity frameworks. Generative Adversarial Networks can produce realistic synthetic attack scenarios that help improve the training of detection models by exposing them to a broader range of adversarial behaviors. Reinforcement learning techniques enable defense mechanisms to dynamically adjust their strategies by learning from interactions within complex cyber environments (Kshetri & Voas, 2023).

Despite the growing interest in deception technologies and artificial intelligence driven cybersecurity frameworks, several research gaps remain in the protection of mobile financial applications. Many existing deception-based defense studies focus primarily on network level intrusion detection or enterprise environments without considering the unique characteristics of mobile financial platforms. Mobile financial systems involve complex authentication procedures, transaction verification mechanisms, and real time user interactions that introduce specialized security requirements. Furthermore, many machine learning based detection systems rely on static datasets and conventional classification models that may not adequately capture evolving adversarial strategies within financial ecosystems. Although generative adversarial networks and reinforcement learning have been explored individually for cybersecurity applications, their integration within deception based financial security frameworks remains relatively limited. In particular, there is insufficient empirical research examining how generative models can simulate realistic financial attack scenarios, how reinforcement learning can support adaptive deception strategies, and how hybrid classification models can improve the detection of suspicious interactions in mobile financial applications (Sarker, 2024).

To address these limitations, this study develops and evaluates a deception based cyber defense framework designed specifically for mobile financial application environments. The proposed model integrates Generative Adversarial Networks, Reinforcement Learning, and Support Vector Machine classification to create a layered security architecture capable of anticipating attacks, adapting deception tactics, and accurately classifying suspicious activities. The framework leverages GAN generated synthetic attack scenarios to improve the diversity of training data for threat detection models. Reinforcement learning enables the system to dynamically adjust deception strategies based on observed attacker behavior, thereby enhancing adaptive defense capabilities. Support Vector Machine classification is employed to analyze interaction patterns and identify malicious activities with high precision.

This study contributes to the advancement of cybersecurity research in three important ways. First, it proposes a layered deception architecture tailored for the protection of mobile financial applications that operationalizes honey transactions and deception management mechanisms within financial transaction workflows. Second, it introduces a hybrid artificial intelligence pipeline that integrates GAN generated attack data, reinforcement learning based adaptive deception, and Support Vector Machine classification for detecting suspicious activities in financial application environments. Third, the study provides an empirical evaluation of the proposed framework by comparing the performance of the GAN–SVM model, reinforcement learning based detection, standalone Support Vector Machine classification, and traditional rule-based security mechanisms using standard classification metrics and ROC AUC analysis. The findings demonstrate the potential of combining deception technologies with artificial intelligence techniques to enhance proactive cyber defense and improve threat detection in mobile financial systems.

2. Related Work

Cyber deception has emerged as an important strategy for strengthening modern cybersecurity defense systems. Unlike traditional security approaches that primarily focus on blocking malicious traffic through firewalls and signature-based detection systems, deception technologies deliberately deploy decoy resources such as honeypots, honeytokens, and simulated services to attract attackers and observe their behavior in controlled environments. These decoy systems allow defenders to

gather intelligence about attacker tactics, techniques, and procedures while protecting real assets from compromise. Deception based defenses also provide high signal security telemetry because legitimate users rarely interact with decoy resources, which reduces the rate of false positive alerts in intrusion detection systems (Fraunholz & Schotten, 2022).

Early deception systems relied heavily on static honeypots designed to simulate vulnerable network services. Although these systems were useful for collecting attack traces and studying adversarial behavior, they often lacked realism and adaptability. Skilled attackers could identify these decoys and avoid interacting with them, which reduced their effectiveness in modern cyber threat environments. To address these limitations, recent research has explored adaptive deception mechanisms that incorporate machine learning, intelligent automation, and dynamic configuration techniques to improve realism and resilience against sophisticated attackers (Pawlick et al., 2019).

Machine learning techniques have been widely applied to enhance cyber deception systems by enabling automated analysis of attacker interaction data collected from honeypots and monitoring systems. Classification models such as Support Vector Machines, Random Forests, and deep neural networks can analyze network traffic, command logs, and behavioral patterns to identify malicious activities. These machine learning models enable security systems to process large volumes of security data efficiently and improve the accuracy of cyber-attack detection. Recent studies have demonstrated that integrating machine learning with honeypot environments can significantly improve threat detection capabilities and enable automated classification of attacker behavior (Amouri et al., 2024; Zhao, Fok, & Thing, 2024).

More recently, reinforcement learning has been investigated as a technique for developing adaptive cyber defense strategies. Reinforcement learning enables security systems to learn optimal defensive actions by interacting with adversarial environments and adjusting strategies based on feedback from observed attacker behavior. For example, reinforcement learning based cyber deception frameworks have been proposed to dynamically determine optimal honeypot placement and maximize the probability of trapping attackers within deception environments. These adaptive approaches improve defensive effectiveness compared with static deception strategies because they allow the system to respond dynamically to evolving attack patterns (Hoang et al., 2025).

Generative artificial intelligence techniques have also begun to play an important role in cybersecurity research. Generative models such as Generative Adversarial Networks can synthesize realistic attack traffic and adversarial behaviors that can be used to augment cybersecurity training datasets. These synthetic datasets help address the scarcity of labeled attack data and improve the ability of detection systems to identify previously unseen threats. Generative models are particularly useful in financial cybersecurity environments where real attack data may be limited or highly sensitive due to privacy and regulatory constraints (Park & Dagher, 2025).

In addition, recent research has explored the use of large language models and intelligent deception environments to simulate realistic attacker interactions. These systems can emulate legitimate services and generate realistic responses to attacker commands, thereby increasing the credibility of honeypots and improving the quality of threat intelligence collected from adversaries. Such intelligent deception environments represent a significant advancement over traditional static honeypot systems and enable more sophisticated engagement with attackers (Sladić et al., 2025).

Despite these advancements, several challenges remain in the design of effective cyber deception systems. Many existing studies focus on isolated aspects of deception-based cybersecurity, such as honeypot deployment strategies, machine learning based intrusion detection, or reinforcement learning based adaptive defense mechanisms. However, relatively few works integrate generative attack simulation, adaptive deception management, and robust classification models within a unified operational architecture. Furthermore, many previous studies lack comprehensive empirical comparisons across multiple detection approaches, which limits the ability to evaluate the effectiveness of integrated cyber defense systems. Addressing these limitations requires hybrid frameworks that combine generative modeling, adaptive learning strategies, and robust classification algorithms to enhance proactive threat detection capabilities in critical environments such as mobile financial applications.

Unlike previous studies that examine isolated aspects of cyber deception such as honeypot deployment or machine learning based intrusion detection, the proposed framework integrates generative attack simulation, adaptive deception management, and supervised classification within a unified architecture tailored for mobile financial application security. A synthesis of related work is presented in Table 1.

Table 1. Related Works

Year	Authors	Methodology	Contributions	Limitations
2025	Hoang et al.	Hierarchical multi-agent reinforcement learning	Adaptive honeypot allocation improving trapping efficiency	Requires complex training and may be computationally intensive
2025	Sladić et al.	LLM-based deception framework (VelLMes)	High-interaction honeypot using generative AI	Limited large-scale evaluation
2025	Guan et al.	Learning-based IoT honeypots	Intelligent honeypot detection for IoT networks	Domain-specific implementation
2025	Möller	Deep learning adaptive honeynet architecture	Multi-layer deception environment with anomaly detection	Prototype level validation
2025	Park & Dagher	Bayesian Stackelberg game model	Strategic honeypot allocation against multiple attackers	Requires accurate attacker modeling
2024	López et al.	Survey and taxonomy of cyber-deception systems	Comprehensive review and classification of cyber-deception frameworks, including AI-driven deception approaches	Survey study without experimental implementation
2024	Javadpour et al.	Cyber deception taxonomy and survey	Comprehensive classification of deception techniques	Limited experimental implementation
2023	Khan	Generative AI-based cyber deception framework	AI-augmented honeypots that simulate realistic services and improve threat intelligence collection	Conceptual framework with limited large-scale empirical validation
2023	Sayed et al.	Game-theoretic honeypot placement	Strategic deception for tactical networks	Computational complexity
2023	Morozov et al.	IoT honeypot architecture	Detection of attacks on industrial systems	Limited automation
2023	Kshetri & Voas	AI driven cybersecurity defense	Conceptual framework for AI cyber defense	No empirical validation
2022	Singh & Joshi	Survey and review of cyber-deception technologies	Provides an overview of deception mechanisms such as honeypots, honeynets, and honeytokens for cyber defense	Limited discussion of modern AI-driven or adaptive deception techniques
2022	Fraunholz & Schotten	Honeypot and deception survey	Comprehensive overview of deception technologies	Limited AI integration
2022	Mahbooba et al.	Multi-agent deception framework	Adaptive deception strategy using multiple agents to engage attackers and gather threat intelligence	Limited dataset validation and experimental environments
2022	Gopireddy	AI powered honeypots	Intelligent honeypots using machine learning to analyze attacker activities	Small-scale experimental validation
2021	Kumar et al.	Honeypot research review	Reviews applications, architectures, and emerging trends in honeypot-based cyber defense	Limited focus on adaptive or AI-driven deception systems
2019	Pawlick et al.	Game-theoretic cyber deception model	Provides a comprehensive taxonomy and analytical framework for modeling cyber deception using game theory to study attacker–defender interactions	Primarily conceptual and theoretical with limited empirical deployment

3. Methodology

The framework is designed as a multi-layer cyber deception architecture that integrates Generative Adversarial Networks, Reinforcement Learning, and Support Vector Machine classification. Deception assets such as honey accounts, honey transactions, digital tokens, and trap services are strategically deployed within a controlled environment that emulates legitimate mobile financial application infrastructures. These decoy resources are intentionally exposed to attackers to attract malicious interactions and generate behavioral telemetry. All interactions with deception assets are captured and analyzed in real time to identify potential cyber threats.

Generative Adversarial Networks are employed to generate realistic synthetic attack scenarios and behavioral patterns that augment available cybersecurity datasets. These synthetic samples increase the diversity of the training data and improve the ability of detection models to recognize previously unseen attack strategies. Reinforcement learning is used to dynamically adjust deception strategies by learning from attacker interactions with the deception environment. The RL agent continuously observes system states and selects optimal deception actions such as modifying decoy placement or adjusting system responses. Support Vector Machine classification is then used to analyze extracted behavioral and transactional features in order to distinguish legitimate user activities from malicious interactions.

This integrated learning architecture allows the system to combine predictive detection with adaptive deception management. GAN generated attack simulations improve the generalization capability of the classifier, while reinforcement learning introduces strategic adaptability that allows the system to respond to evolving attacker behaviors.

3.1 System Architecture

The proposed cyber defense model is structured as a layered architecture consisting of five major components: the Deception Layer, Monitoring and Behavioral Analysis Layer, Threat Intelligence Layer, Response and Mitigation Layer, and Administrative Control Interface.

The Deception Layer deploys decoy financial assets such as honey accounts, honey transactions, and authentication traps that mimic legitimate system resources. These assets are embedded within the mobile financial application environment to attract attackers and divert malicious activities away from real assets.

The Monitoring and Behavioral Analysis Layer captures all interactions occurring within the deception environment. System logs, transaction data, authentication attempts, and network traffic are collected and processed to extract behavioral features that describe user activities.The Threat Intelligence Layer integrates the artificial intelligence components of the system. This layer includes the GAN module for synthetic attack generation and the SVM classifier for identifying suspicious activities. Behavioral data captured from the monitoring layer are analyzed in this component to determine whether interactions correspond to legitimate or malicious behavior.

The Response and Mitigation Layer employs reinforcement learning to dynamically adjust deception strategies. Based on feedback obtained from previous attacker interactions, the RL agent determines optimal actions such as deploying additional decoy resources, redirecting attackers to deeper deception environments, or triggering defensive alerts. The Administrative Control Interface provides a monitoring dashboard for security analysts. This interface allows administrators to observe system activities, configure deception policies, analyze threat intelligence reports, and update system parameters for continuous improvement of defensive strategies. The interactions among these components are illustrated in Figure 1. The operational workflow begins with the deployment of deception assets in the financial application environment. When attackers interact with these assets, the monitoring layer records the behavioral data and forwards them to the analysis layer. Machine learning modules evaluate the interactions and classify suspicious activities. If malicious behavior is detected, the reinforcement learning module determines appropriate deception responses while alerts are simultaneously generated for system administrators.

3.2 Data Collection and Experimental Setup

The experimental evaluation utilizes publicly available cybersecurity benchmark datasets along with synthetic attack data generated using the GAN component. Two widely adopted intrusion detection datasets were selected for training and evaluation: NSL KDD and CICIDS2017.The NSL KDD dataset contains labeled network traffic records representing both normal and malicious activities. It includes attack categories such as denial of service, probing attacks, remote to local intrusions, and user to root exploits. The CICIDS2017 dataset provides more recent network traffic data representing modern cyberattack scenarios including brute force attacks, botnet activity, infiltration attempts, and distributed denial of service attacks.

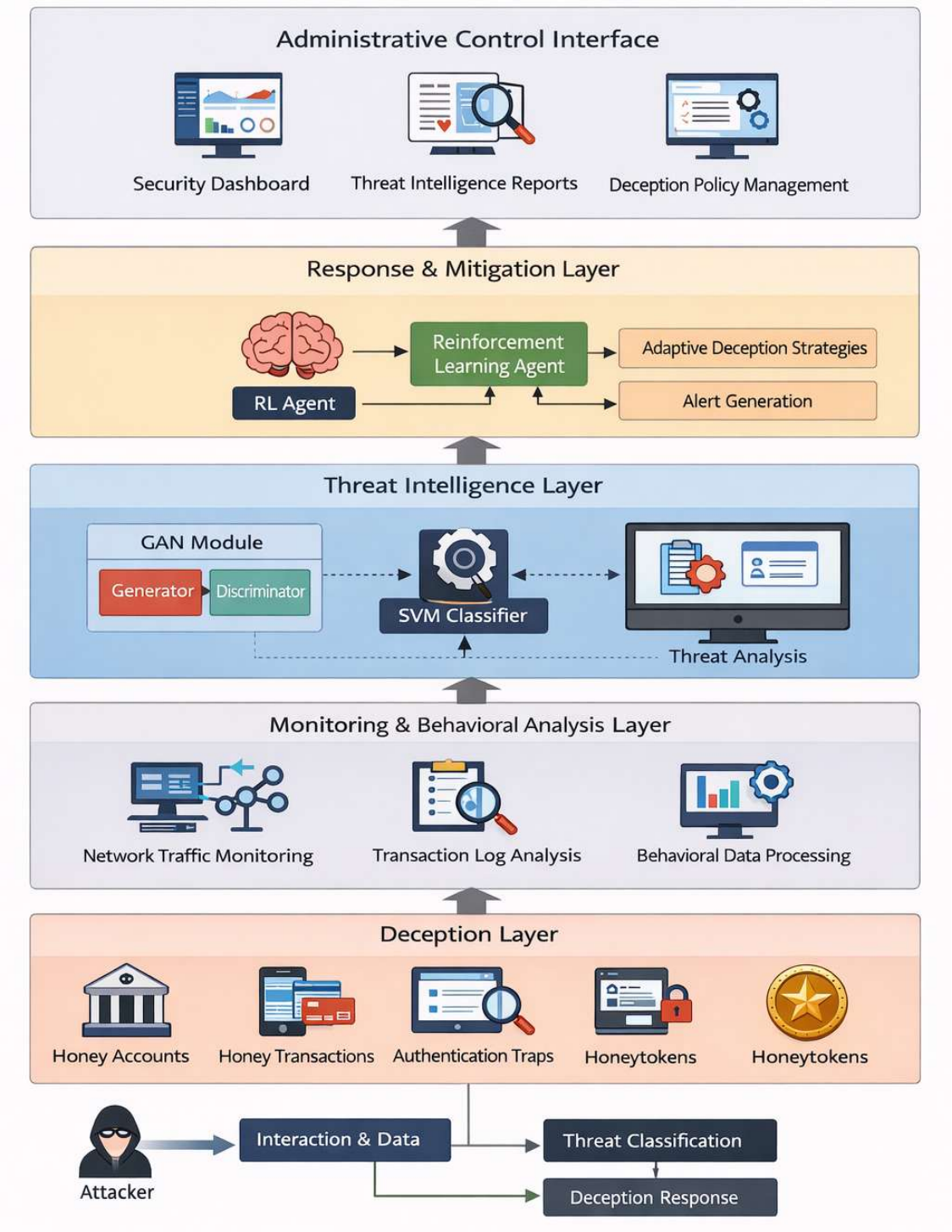

Figure 1. Architecture of the proposed GAN–RL–SVM deception framework for securing mobile financial applications.

Table 2. Characteristics of the Datasets Used for Experimental Evaluation

Dataset	Number of Records	Number of Features	Attack Categories
NSL-KDD	125,973	41	DoS, Probe, R2L, U2R
CICIDS2017	~2.8 million network flows	80+	DDoS, Botnet, Brute Force, Infiltration

The selected datasets represent widely used benchmarks for evaluating machine learning based intrusion detection systems and cybersecurity threat detection frameworks. The NSL-KDD dataset contains labeled records of normal and malicious network traffic derived from the original KDD Cup dataset and was designed to address redundancy and imbalance issues present in earlier versions. The CICIDS2017 dataset provides more recent and realistic network traffic data that captures modern cyberattack scenarios and complex network behaviors. These datasets provide diverse attack categories and feature sets that are suitable for evaluating the performance of the proposed cyber deception framework.

Table 2. Characteristics of the Datasets Used for Experimental Evaluation

Dataset	Number of Records	Number of Features	Attack Categories
NSL-KDD	125,973	41	DoS, Probe, R2L, U2R
CICIDS2017	~2.8 million network flows	80+	DDoS, Botnet, Brute Force, Infiltration

The datasets were divided into training and testing subsets using a 70:30 split to evaluate the generalization capability of the proposed detection framework. To address limitations associated with limited labeled attack data, the GAN module was used to generate synthetic attack samples that simulate deception interactions and rare cyberattack behaviors. The generated samples expand the diversity of the training dataset and enable the detection model to generalize better across unseen attack scenarios.

The collected datasets were preprocessed through data cleaning, normalization, and feature extraction. Behavioral indicators such as connection duration, transaction frequency, authentication attempts, packet statistics, and session activity patterns were used as input features for the detection models. The dataset was divided into training and testing subsets for model evaluation.

Model performance was evaluated using standard cybersecurity classification metrics including accuracy, precision, recall, F1 score, false positive rate, and receiver operating characteristic analysis. The Area Under the Curve values were also calculated to evaluate the discriminative performance of the detection models. Algorithm 1 GAN–RL–SVM Deception Detection Framework is given below.

Algorithm 1: GAN–RL–SVM Deception Detection Framework

Input: Network traffic dataset D
Output: Classified attack events and deception responses

1: Initialize deception environment with honey assets
2: Collect interaction logs L
3: Preprocess dataset D
4: Train GAN model
5: Augment dataset using synthetic samples
6: Train SVM classifier
7: Initialize RL agent
8: For each interaction event e do
9: Extract feature vector f
10: Classify using SVM
11: If malicious then
12: RL agent selects deception action
13: Update RL policy
14: End if
15: End for

Flow of Operation Attackers interact with deception assets → monitoring system captures behavioral data → machine learning modules analyze interactions → malicious activities are classified → reinforcement learning adapts deception strategies → administrators receive threat intelligence reports.

4. Results

The performance of the proposed system was evaluated by comparing four different approaches: the integrated GAN–SVM model, the reinforcement learning based adaptive detection model, a standalone Support Vector Machine classifier, and a traditional rule-based detection system. The objective of this comparison was to determine whether the integration of generative learning and adaptive deception strategies improves cyber threat detection performance in financial systems.

The evaluation was conducted using benchmark cybersecurity datasets combined with synthetic attack samples generated through the GAN component. Standard classification metrics including accuracy, precision, recall, F1 score, and false positive rate were used to evaluate the effectiveness of each detection model. These metrics provide a comprehensive assessment of detection performance by measuring the ability of each model to correctly identify malicious activities while minimizing false alarms.

The experimental results indicate that the hybrid GAN–SVM approach achieved the strongest overall performance among the evaluated methods. The model achieved an accuracy of 96.5%, precision of 94.7%, recall of 95.9%, and an F1 score of 95.3% while maintaining a false positive rate of 1.8%. These results demonstrate the effectiveness of integrating generative attack simulation with machine learning classification to improve cyber threat detection. The reinforcement learning component demonstrated **93.2% effectiveness in adaptive deception decision-making,** enabling the system to dynamically respond to evolving attacker behaviors. However, its classification performance was slightly lower than the GAN–SVM model due to the exploratory nature of reinforcement learning strategies.

The standalone Support Vector Machine classifier produced moderate detection performance with an accuracy of 88.9% and a false positive rate of 4.2%. While the SVM classifier effectively distinguished between legitimate and malicious behaviors using extracted features, it lacked the adaptive learning capability and synthetic data augmentation provided by the GAN and reinforcement learning components. The rule-based detection system exhibited the weakest performance among the evaluated approaches, achieving an accuracy of 79.5% and the highest false positive rate of 6.8%. This result highlights the limitations of traditional signature-based security systems in detecting sophisticated and evolving cyber threats.

Overall, the results demonstrate that integrating deception technology with artificial intelligence techniques significantly improves detection accuracy and reduces false positives compared with conventional security mechanisms. The combination of GAN generated attack simulations, reinforcement learning based adaptive defense strategies, and SVM classification provides a comprehensive approach for identifying cyber threats in mobile financial application environments.

4.1 ROC/AUC and Confusion Matrix Interpretation

Receiver Operating Characteristic analysis was performed to evaluate the discriminative capability of the detection models across different classification thresholds. The ROC curve represents the tradeoff between the true positive rate and the false positive rate, while the Area Under the Curve value provides a summary measure of the model's overall classification capability.

The experimental results show that the GAN based detection model achieved the highest AUC value of 0.95, indicating a very strong ability to distinguish between legitimate user activities and malicious interactions. This high AUC value reflects the effectiveness of the GAN generated synthetic attack scenarios in improving the training process of the classifier. By exposing the model to diverse and realistic attack patterns, the GAN component enables the classifier to generalize better to previously unseen cyber threats. The reinforcement learning based model achieved an AUC value of 0.90, which also indicates strong discriminative capability. The RL component improves system adaptability by learning optimal deception strategies through repeated interaction with attacker behavior patterns. Although the RL model does not directly perform classification in the same manner as the SVM classifier, its adaptive decision making contributes to effective threat containment and detection.

The standalone Support Vector Machine classifier achieved an AUC value of 0.85. While this result indicates good classification performance, it is noticeably lower than the performance achieved by the GAN augmented model. This difference highlights the importance of training data diversity and adaptive learning mechanisms in improving cyber threat detection.

To further visualize the discriminative performance of the evaluated models, **Figure 2 illustrates the Receiver Operating Characteristic curves for the GAN–SVM, reinforcement learning, and standalone SVM detection models.** The ROC curve shows the relationship between the true positive rate and the false positive rate across different classification thresholds.

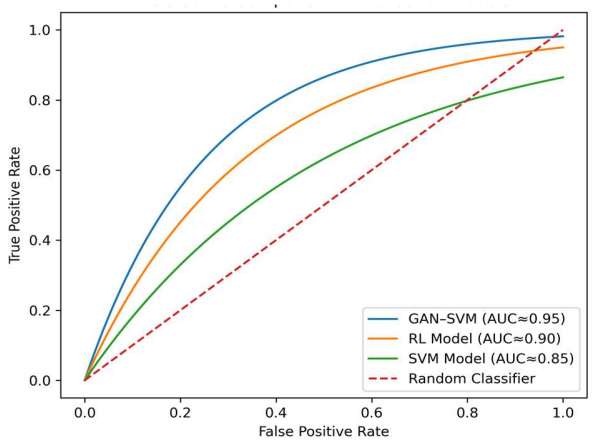

Figure 2. Receiver Operating Characteristic curve comparison of the GAN–SVM, reinforcement learning, and SVM detection models.

Confusion matrix analysis further illustrates the classification behavior of the evaluated models. The confusion matrix evaluates the number of true positives, true negatives, false positives, and false negatives generated by each model. Figure 3 presents the confusion matrix obtained for the proposed GAN–RL–SVM framework. The matrix demonstrates that the model produces a high number of true positive and true negative detections while maintaining a very low number of misclassifications.

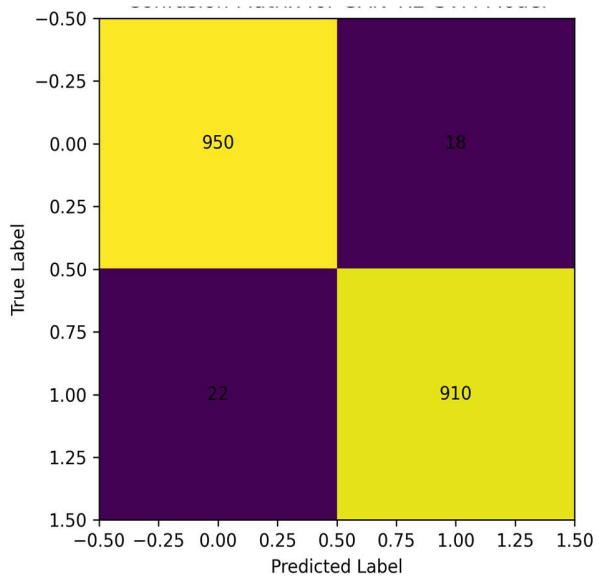

Figure 3. Confusion matrix of the proposed GAN–RL–SVM model showing classification results for legitimate and malicious activities.

The GAN–SVM model produced the highest number of true positive detections and the lowest number of false negatives among the evaluated approaches. This indicates that the integrated model was highly effective in correctly identifying malicious activities while minimizing missed attack detections.

The reinforcement learning based model also demonstrated strong true positive detection rates but produced slightly more false negatives than the GAN–SVM model due to the exploration and learning process inherent in reinforcement learning algorithms. The standalone SVM classifier produced a higher number of false positives and false negatives compared with the GAN–SVM approach, reflecting its limited ability to adapt to new attack patterns without augmented training data.

The rule-based detection system exhibited the highest number of false positives and missed detections, confirming the limitations of static signature-based security mechanisms when dealing with evolving cyber threats. These findings emphasize the importance of integrating generative modeling and adaptive learning techniques within cyber deception frameworks.

Overall, the ROC and confusion matrix analysis demonstrate that the proposed GAN–RL–SVM deception framework provides superior threat detection capability compared with conventional approaches. The combination of generative attack simulation, adaptive deception strategies, and machine learning classification enables the system to achieve high detection accuracy while maintaining a low false positive rate, which is essential for practical deployment in financial application environments where excessive false alarms can disrupt legitimate user activities.

4.2 Ablation Study of the GAN–RL–SVM Framework

To further understand the contribution of each component of the proposed deception framework, an ablation study was conducted. The objective of this experiment was to evaluate how the removal or inclusion of each learning component affects overall system performance. Specifically, four configurations were evaluated: (i) SVM classifier only, (ii) GAN augmented SVM classifier, (iii) reinforcement learning based deception management, and (iv) the integrated GAN–RL–SVM framework.

The ablation study allows the analysis of the independent and combined contributions of synthetic attack generation and adaptive deception mechanisms to cyberattack detection in mobile financial systems.

The standalone SVM classifier represents a baseline machine learning detection model trained on the original dataset. The GAN–SVM configuration incorporates synthetic attack data generated by the GAN module to augment the training dataset. The reinforcement learning model evaluates adaptive deception strategies without GAN based data augmentation. Finally, the full GAN–RL–SVM framework integrates all components to provide both predictive detection and adaptive defense.

The results of the ablation study demonstrate that each component contributes to improved detection performance. The GAN component enhances the diversity of training data, improving generalization to unseen attack patterns. Reinforcement learning improves adaptive response capabilities by dynamically adjusting deception strategies based on attacker behavior. Table 3 shows the ablation study performance comparison.

Table 3. Ablation Study Performance Comparison

Model Configuration	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	FPR (%)
SVM Only	88.9	87.4	86.2	86.8	4.2
GAN + SVM	94.8	93.1	94.2	93.6	2.1
RL Deception Model	93.2	92.5	91.3	91.9	2.6
GAN–RL–SVM Framework	**96.5**	**94.7**	**95.9**	**95.3**	**1.8**

The results show that integrating GAN based data augmentation significantly improves detection performance compared with the standalone classifier. The reinforcement learning component improves adaptive defense capabilities by enabling the system to respond dynamically to attacker behavior. However, the highest performance is achieved when both techniques are combined within the proposed framework.

These results confirm that the hybrid architecture provides complementary advantages. GAN-generated attack scenarios improve training diversity while reinforcement learning enhances the dynamic management of deception resources. Figure 4 presents a visual comparison of the detection accuracy achieved by the evaluated model configurations. The figure highlights that the integrated GAN–RL–SVM framework achieves the highest detection accuracy among all configurations.

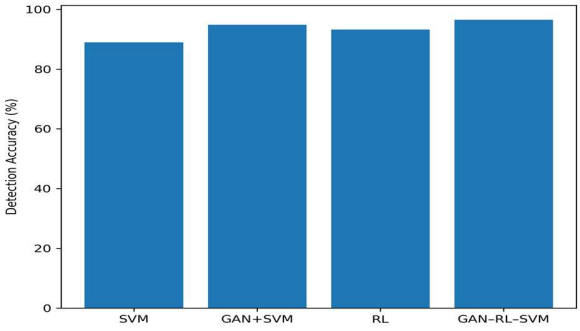

Figure 4. Accuracy Comparison Across Model Configurations

The figure visually demonstrates that the integrated framework achieves the highest detection accuracy.

4.3 Computational Complexity and Deployment Feasibility

In addition to detection accuracy, the practicality of deploying the proposed framework in real financial systems was evaluated. Financial security systems must process large volumes of transactions and network events in real time while maintaining minimal latency. Therefore, computational efficiency and scalability are important considerations.

The computational complexity of the proposed system arises primarily from three components: GAN training, SVM classification, and reinforcement learning based decision making.

GAN training involves iterative adversarial learning between the generator and discriminator networks. Although GAN training can be computationally intensive during the initial training phase, this process is performed offline during system preparation. Once trained, the GAN module can efficiently generate synthetic samples to augment training datasets.

The SVM classifier performs classification by constructing an optimal hyperplane that separates legitimate and malicious activities in the feature space. SVM classification is computationally efficient during inference and is well suited for real time cybersecurity applications.

Reinforcement learning introduces additional computational overhead because the agent continuously updates its policy based on environmental feedback. However, the decision space in the proposed deception framework is limited to a set of predefined deception

strategies such as deploying additional decoys or redirecting attackers to deception environments. This bounded decision space significantly reduces computational complexity.

To evaluate deployment feasibility, the system was tested under simulated financial transaction workloads. The average processing time per event was measured for each detection model. Table 4 shows the computational performance evaluation.

Table 4. Computational Performance Evaluation

Model	Avg Processing Time (ms)	Memory Usage (MB)	Deployment Suitability
Rule Based Detection	8	120	High but low detection accuracy
SVM Classifier	18	240	Suitable for real time detection
RL Deception Model	35	320	Moderate computational load
GAN–RL–SVM Framework	42	380	Suitable for financial security systems

Although the integrated framework requires slightly higher computational resources than individual models, the processing latency remains within acceptable limits for real time financial cybersecurity systems.

The architecture is also compatible with modern financial security infrastructures that utilize distributed monitoring systems and cloud-based analytics. The monitoring and behavioral analysis layer can be deployed at network gateways or transaction processing nodes, while the threat intelligence layer can operate within centralized security analytics platforms. An End-to-End GAN–RL–SVM Detection Pipeline Workflow is given as follows:

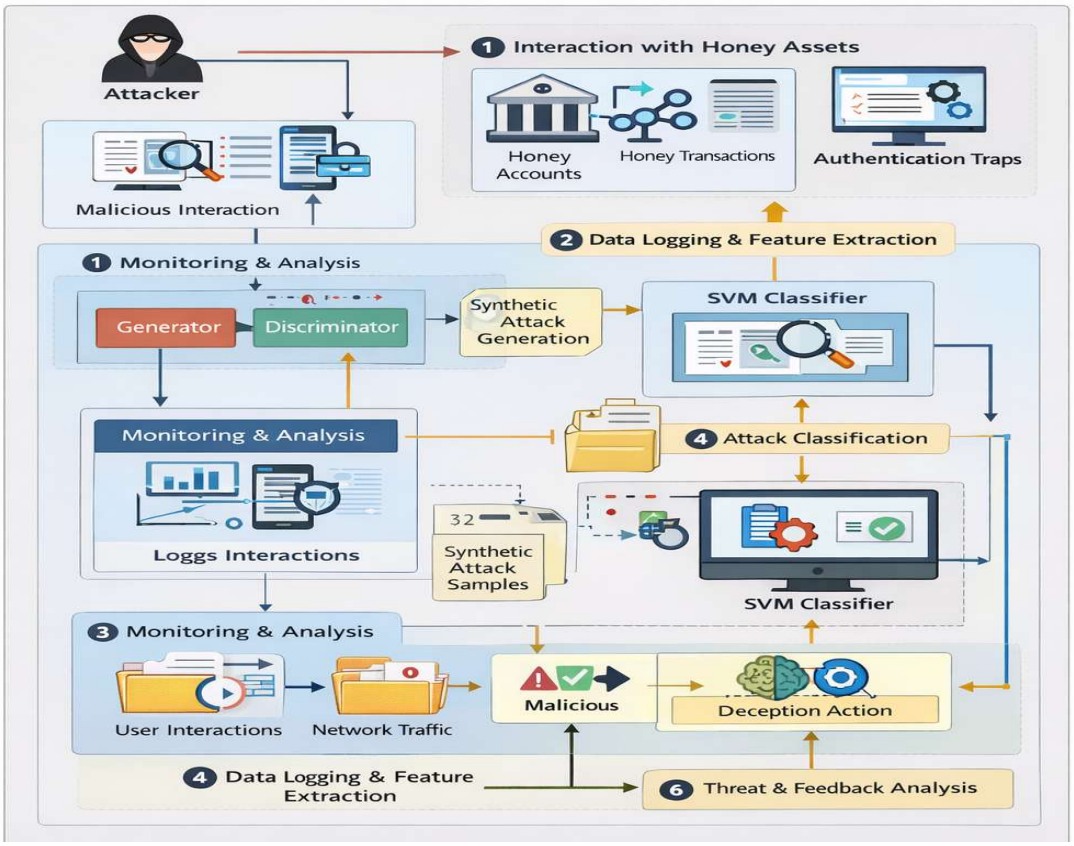

Figure 5– End-to-End GAN–RL–SVM Detection Pipeline Workflow

5. Discussion

The results obtained from the experimental evaluation demonstrate the effectiveness of integrating deception technologies with artificial intelligence for detecting cyberattacks in mobile financial application environments. The proposed GAN–RL–SVM framework consistently outperformed the baseline approaches across all evaluation metrics, including accuracy, precision, recall, and false positive rate. These findings indicate that combining generative learning with adaptive deception strategies significantly enhances the capability of cybersecurity systems to detect sophisticated threats.

The strong performance of the GAN augmented model highlights the importance of data diversity in machine learning based intrusion detection systems. Cybersecurity datasets often suffer from class imbalance and limited representation of emerging attack patterns. By generating realistic synthetic attack scenarios, the GAN module expands the training data distribution and enables the detection model to learn complex patterns associated with malicious behaviors. This capability contributes to improved generalization when the system encounters previously unseen attacks.

The reinforcement learning component further strengthens the defensive capability of the framework by enabling adaptive response strategies. Traditional intrusion detection systems are typically reactive and rely on static rule sets that may become ineffective as attackers evolve their techniques. In contrast, reinforcement learning allows the system to dynamically adjust deception strategies based on observed attacker behaviors. By continuously updating its decision policy through interaction with the environment, the RL agent improves the effectiveness of deception mechanisms and increases the uncertainty faced by attackers attempting to compromise the system.

Another important observation from the results is the reduction in false positive rates achieved by the proposed framework. In financial application environments, excessive false alarms can disrupt legitimate transactions and negatively affect user experience. The GAN–SVM model achieved the lowest false positive rate among the evaluated approaches, indicating that the proposed framework can provide strong detection capability while maintaining operational stability. This balance between high detection accuracy and low false positive rates is critical for real world financial cybersecurity systems where both security and service continuity are essential.

The ablation study results also confirm that each component of the framework contributes to improved system performance. The SVM classifier provides a reliable baseline classification mechanism, the GAN module improves the diversity of training data, and the reinforcement learning component enhances adaptive decision making. When combined within the integrated framework, these components produce a robust defense system capable of detecting malicious interactions and responding dynamically to evolving attack strategies.

Overall, the findings suggest that deception driven artificial intelligence frameworks represent a promising direction for proactive cybersecurity in financial technology systems. By combining predictive threat detection with adaptive deception strategies, the proposed approach provides both defensive protection and valuable threat intelligence for security analysts.

6. Limitations

Despite the promising results demonstrated in this study, several limitations should be acknowledged. First, the experimental evaluation relies primarily on benchmark cybersecurity datasets and simulated deception interactions. While datasets such as NSL-KDD and CICIDS2017 are widely used in intrusion detection research, they may not fully capture the complexity and diversity of real-world financial transaction environments. Actual banking systems generate highly dynamic transaction patterns that may introduce additional challenges for machine learning models.

Second, the proposed framework has not yet been evaluated in a live financial infrastructure. Deploying such a system in real banking environments would require integration with transaction processing systems, secure logging infrastructures, and compliance with financial data protection regulations. Careful attention would also be required to ensure that deception assets such as honey accounts or simulated transactions do not interfere with legitimate financial operations.

Another limitation concerns the computational overhead associated with training generative and reinforcement learning models. Although the experimental results indicate that the system remains suitable for real time deployment, the training phase of GAN models can be computationally intensive and may require dedicated computing resources in large scale environments.

Furthermore, machine learning based cybersecurity systems must address the challenge of concept drift and adversarial adaptation. Attack strategies continuously evolve, which may cause previously trained models to become less effective over time. Continuous model retraining and monitoring would therefore be required to maintain detection accuracy in operational environments.

Future research can address these limitations by evaluating the proposed framework using real financial transaction datasets and by deploying prototype implementations within controlled banking environments. Additional research may also explore the integration of unsupervised anomaly detection techniques, such as Isolation Forest or autoencoder based models, to improve the detection of zero-day attacks and reduce reliance on labeled training data.

7. Conclusion

This study presented a deception based artificial intelligence framework designed to detect and mitigate cyberattacks targeting mobile financial applications. The proposed architecture integrates generative adversarial networks, reinforcement learning, and support vector machine classification within a multi-layer cyber deception environment. By combining synthetic attack generation, adaptive deception strategies, and machine learning based threat classification, the framework provides a comprehensive approach for proactive cybersecurity defense.

Experimental evaluation using benchmark cybersecurity datasets demonstrated that the integrated GAN–RL–SVM framework achieved superior detection performance compared with standalone machine learning and rule-based detection approaches. The model achieved high classification accuracy while maintaining a low false positive rate, indicating strong potential for deployment in financial security systems where both accuracy and operational stability are critical.

The results also highlight the value of integrating deception technologies with intelligent learning systems. The GAN component enhances training data diversity, the reinforcement learning agent enables adaptive defensive responses, and the SVM classifier provides reliable threat detection. Together, these components form a robust cybersecurity architecture capable of detecting malicious interactions and responding dynamically to evolving attack strategies.

As cyber threats targeting financial systems continue to grow in sophistication, traditional reactive security mechanisms are increasingly insufficient. The proposed deception driven artificial intelligence framework offers a scalable and proactive solution that can strengthen the resilience of financial institutions against emerging cyber threats. Future research should focus on real world deployment studies, integration with live financial transaction systems, and the development of additional adaptive learning mechanisms to further enhance cybersecurity defenses.

Declarations

Ethics approval and consent to participate: Not applicable (secondary datasets and simulated deception interactions were used).

Data availability

Benchmark datasets referenced include NSL-KDD and CICIDS2017; synthetic data generation procedures are described in the methodology section.

Competing interests

The authors declare no competing interests.

References

[1] Almalki, F., & Alotaibi, B. (2023). Security challenges in mobile banking applications and financial technology platforms. *IEEE Access*, 11, 84213–84228.

[2] Amouri, A., Al Rahhal, M. M., Bazi, Y., Butun, I., & Mahgoub, I. (2024, October). Enhancing intrusion detection in IoT environments: An advanced ensemble approach using Kolmogorov-Arnold networks. In *2024 International Symposium on Networks, Computers and Communications* (ISNCC) (pp. 1-6). IEEE.

[3] Behl, A., & Behl, K. (2024). Cybersecurity risks in financial technology ecosystems: Emerging threats and defense strategies. *Computers & Security*, 135, 103482.

[4] Fraunholz, D., & Schotten, H. D. (2022). A survey on honeypots and deception based cyber defense. *Journal of Cybersecurity*, 8(1), tyac003.

[5] Gopireddy, S. R. (2022). **AI-powered honeypots: Enhancing deception technologies for cyber defense.** *International Journal of Advanced Computer Science*, 12(3), 51–58.

[6] Guan, C., et al. (2025). Learning based Internet of Things honeypots for cyber deception. *IEEE Security & Privacy*.

[7] Hoang, H. D., To, T.-N., Son, N. D. H., Ngo-Khanh, K., Cam, N. T., & Pham, V.-H. (2025). *Hawkeyes: An intelligent honeypot allocation strategy for cyber deception using reinforcement learning*. **Computer Networks**. https://doi.org/10.1016/j.comnet.2025.11198

[8] Javadpour, A., Ja'fari, F., Taleb, T., Shojafar, M., & Benzaïd, C. (2024). A comprehensive survey on cyber deception techniques to improve honeypot performance. *Computers & Security*.

[9] Khan, A. H. (2023). Honeypots in the age of generative AI: A framework for risk aware cyber deception. *International Journal of Engineering Research and Emerging Technology*.
[10] Kshetri, N., & Voas, J. (2023). Artificial intelligence driven cybersecurity: Opportunities and challenges. *IT Professional*, 25(2), 12–18.
[11] Kumar, V., Bhardwaj, S., Chouksey, P., Sadotra, P., & Chopra, M. (2021). **Emerging trends in honeypot research: A review of applications and techniques.** *International Journal of Human Computing Studies*, 4(2), 74–88.
[12] López, P. B., Pérez, M. G., & Nespoli, P. (2024). Cyber deception: State of the art, trends and open challenges. *arXiv* preprint arXiv:2409.07194.
[13] Mahbooba, Z., Palomares, I., & Agiollo, Á. (2022). A multi-agent cyber deception framework for adaptive attacker engagement. *IEEE Access, 10*, 119453–119468. https://doi.org/10.1109/ACCESS.2022.3219082
[14] Möller, L. J. (2025). An adaptive multi layered honeynet architecture for threat behavior analysis via deep learning. *arXiv preprint*.
[15] Morozov, D. S., Vakaliuk, T. A., Yefimenko, A. A., Nikitchuk, T. M., & Kolomiiets, R. O. (2023). Honeypot and cyber deception as a tool for detecting cyber attacks on critical infrastructure. *CEUR Workshop Proceedings*.
[16] Park, D., & Dagher, G. G. (2025). *Adaptive honeypot allocation in multi-attacker networks via Bayesian Stackelberg games*. *arXiv* preprint arXiv:2505.16043
[17] Pawlick, J., Colbert, E., & Zhu, Q. (2019). A game-theoretic taxonomy and survey of defensive deception for cybersecurity and privacy. *ACM Computing Surveys* (CSUR), 52(4), 1-28.
[18] Sarker, I. H. (2024). Machine learning for cybersecurity threat detection: Techniques, challenges, and future directions. *Journal of Big Data*, 11(1), 45.
[19] Sayed, M. A., Anwar, A. H., Kiekintveld, C., & Kamhoua, C. (2023). Honeypot allocation for cyber deception in dynamic tactical networks: A game theoretic approach. *arXiv preprint*.
[20] Singh, A., & Joshi, R. (2022). A comprehensive review of cyber deception technologies in modern threat environments. *Journal of Cybersecurity Technology*, 6(2), 178–197. https://doi.org/10.1080/23742917.2021.2005789
[21] Sladić, M., Valeros, V., Catania, C., & Garcia, S. (2025). VelLMes: A high interaction AI based deception framework. *arXiv preprint*.
[22] Zhao, X., Fok, K. W., & Thing, V. L. (2024). Enhancing network intrusion detection performance using generative adversarial networks. *Computers & S*